Real-time signal estimation: an efficient alternative to traditional model-based approaches

Marc Wildi
Real-time detection of turning-points for the new KOF-Konjunkturbarometer

Marc Wildi
Institute for Data Analysis and Processdesign (idp), ZHW, Zurich,
www.idp.zhwin.ch
Institute for Business Cycle Research (KOF), ETH, Zurich,
www.kof.ethz.ch
Email: wia@zhwin.ch

The KOF-Konjunkturbarometer is an important leading indicator for Swiss GDP, see http://www.kof.ethz.ch/deutsch/konjunktur/konjunktur_barometer.html. The original indicator is now twenty years old and during this period it lost part of its lead on the GDP series. The new indicator, released on April 2006, is based on an extended data set and on a new `real time` filter technique, the so called Direct Filter Approach (DFA) proposed in Wildi [4]. We here introduce informally the main concepts behind the filtering method.

`Smoothing` is necessary because the series defining the indicator are subject to short-term (infra-year) variations due to `noise` and seasonal components which should not be assigned to the business cycle. Therefore, the undesirable components must be eliminated by designing a suitable filter, see (1.3) below. We call the resulting smoothed component `ideal trend`. Turning-points (TP’s) of the Konjunkturbarometer are defined by the local extrema of this component.

Figure 1 shows the unfiltered aggregate (indicator) and the ideal trend. One can notice that local extrema of the trend and of the original series often disagree: this effect is caused by the noisy short-term components which affect the location of the TP’s. As can be seen, a TP of the smooth trend component is generally strongly indicative for a change in the dynamic of the original series. Therefore, a statistical procedure which would detect TP’s of the ideal trend (almost) in real-time would be of great practical interest for the prospective analysis of the time series.
Unfiltered and filtered Konjunkturbarometer

Figure 1: ideal trend

One can observe in the above figure that trend values are missing towards the boundaries of the time series. The missing values are a consequence of the ‘symmetry’ of the ideal trend which is defined in each time point \( t \) as a weighted average of past, contemporary and future (relative to \( t \)) observations

\[
T_t = b_{-k} X_{t-k} + b_{-(k-1)} X_{t-(k-1)} + \ldots + b_{k-1} X_{t+(k-1)} + b_k X_{t+k}
\]

where \( b_{-k} = b_k \) (symmetry) und \( \sum_{j=-k}^k b_j = 1 \). More precisely, the ideal trend \( T_t \) ist the output of a symmetric filter whose weights are determined by the requirement, that short-term (infra-year) components should be eliminated. In the frequency domain, the (real) transfer function of the filter is

\[
\Gamma(\omega) = \begin{cases} 
1, & 0 \leq |\omega| \leq \pi/9 \\
\frac{\pi/7 - |\omega|}{\pi/7 - \pi/9}, & \pi/9 \leq \omega \leq \pi/7 \\
0, & \pi/7 \leq \omega \leq \pi
\end{cases}
\]

from which the filter coefficients
are derived\textsuperscript{1}. The weights of the filter never vanish but they converge to 0 at the rate $1/k^2$ which implies that the above bi-infinite symmetric ideal trend can be closely approximated by a truncated finite trend, for which $\tilde{b}_k = 0, k > 30$. It is this particular truncated and renormalized ($\sum_{|k| \leq 20} \tilde{b}_k = 1$) trend definition which is chosen as our signal, i.e.

$$
(1.4) \quad \tilde{b}_k = \sum_{|j| \leq 30} b_j, \quad |k| \leq 30
$$

$$
0, \text{ otherwise}
$$

Evidently, the symmetric filter cannot be used for detecting TP's in real-time i.e. at the current boundary because of missing values. Therefore, the ideal trend has to be suitably approximated by an asymmetric filter, the so-called concurrent filter. The corresponding estimation problem is our main contribution and is presented informally in the following sections.

Until now the KOF-Konjunkturbarometer was smoothed by the well known X-11 procedure. Therefore, the latter is our benchmark in the following comparisons. Empirical results in Stier/Schips [2] show that the transfer function of the X-11 trend often exhibits non-negligible side-lobs in the stop band, after the fundamental seasonal frequency $\pi/6$, so that undesirable infra-year components may alter the dating of the turning-points, as can be seen in the following figures.

\textsuperscript{1} The smooth transition between $\pi/9$ and $\pi/7$ in (1.2) ensures that the filter weights $b_k$ decay sufficiently rapidly to 0. Also, one can easily verify that $b_{-k} = b_k$. 

The leakage of the X-11 trend transfer function in the stop band causes the observed shifts of the TP’s. The ideal trend is preferred here because it is easier to interpret, because all undesirable short term components are eliminated and because it allows for a formal definition (1.4) which is necessary for stating the estimation problem properly\(^2\). Despite some more or less obvious disadvantages, the X-11 trend signal is used in our empirical comparisons because it has become an established ‘standard’. Consequently, performance measures will be related to both reference signals.

**Estimation problem**

**Efficient level filter: direct filter approach (DFA)**

Assume \(Y_i, \hat{Y}_i\) are the outputs of two filters with transfer functions \(\Gamma(\omega), \hat{\Gamma}(\omega)\), where items associated to the concurrent filter are denoted by a ‘hat’. The parameters of the concurrent filter should be determined by minimizing the mean-square filter error \(E[(Y_i - \hat{Y}_i)^2]\). Although this expression is generally unknown, an efficient estimate is available. For a stationary process\(^3\) \(X_t\), Wildi [3] shows that

\(^2\) The X-11 trend is defined implicitly and a formal definition as in (1.4) is lacking.

\(^3\) Generalizations to non-stationary processes are provided in Wildi [3]. In the particular context of the above leading indicator series are bounded and therefore integrated processes would be a misspecification, see Wildi/Schips [4]. As a result, we here consider the case of stationary processes only.
(1.5) \[ \frac{2\pi}{T} \sum_{k=0}^{[T/2]} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 I_{NX}(\omega_k) = \frac{1}{T} \sum_{t=1}^{T} (Y_t - \hat{Y}_t)^2 \approx E[(Y_t - \hat{Y}_t)^2] \]

where

- \( I_{NX}(\omega_k) \) is the periodogram of the input signal \( X_t \) and \( \omega_k = \frac{k2\pi}{T} \).

- The error in the first approximation in (1.5) is of smaller magnitude \( o(1/\sqrt{T}) \) than that in the second one (superconsistency).

- The arithmetic mean of the squared filter errors is an asymptotically efficient estimate of the unknown mean-square filter error, which is to minimize.

Therefore, the frequency-domain estimate on the left hand-side of (1.5) is an asymptotically efficient estimate of the unknown mean-square filter error\(^4\). Since one can show that these results are valid uniformly (see Wildi [3]), the output \( \hat{Y}_t \) of the filter solving the minimization problem

\[
(1.6) \quad \min_{\hat{\Gamma}} \frac{2\pi}{T} \sum_{k=0}^{[T/2]} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 I_{NX}(\omega_k)
\]

is an (asymptotically) efficient estimate of the (generally unknown) output \( Y_t \) of the symmetric filter, see Wildi [3]. The optimization criterion (1.6) allows for a straightforward interpretation: symmetric and asymmetric filters should match closely at frequencies for which the input signal is dominated by strong components.

The resulting real-time signal estimation procedure is called direct filter approach (DFA) in Wildi [3]. An extensive empirical study based on business survey data in Wildi/Schips [4] shows that the DFA strongly outperforms traditional model-based real-time estimates based on X-12-ARIMA or TRAMO/SEATS. Gains between 20\% and 40\% (reduction of filter error variance) are obtained in the mean depending on sample lengths or data sets (more than 100 series were considered). The important gain obtained by the DFA is a consequence of its efficiency but also of the methodological problems associated with model-based procedures, see Wildi/Schips [4] for details.

Naturally, important characteristics of the concurrent filter such as summarized in the amplitude (selectivity) and the time delay functions deviate from those of the symmetric filter as shown in the following figure.

\(^4\) Note that the first term can be computed whereas the others are generally unknown.
In particular, there is a noticeable leakage in the stop band of the concurrent filter. Therefore, we analyze if this filter also performs well for detecting TP’s of the ideal trend in real time.

**Real-time detection of TP’s**

As can be seen in the following Figure 4, the best concurrent level filter (dotted line) is often insufficiently informative about the occurrence of TP’s.
Whereas the TP took place in *July* 2004, an approximately confident diagnostic about its occurrence based on the concurrent level filter is possible in *October* 2004 only (as we shall see, X-11 performs even worse). The reason for the above phenomenon is related to the flatness of the trend in the vicinity of TP’s which implies that the noisy high frequency components resulting from the leakage of the concurrent filter in the stop band (see Figure 3) become visible. Of all time points this undesirable leakage-effect distorts the signal mostly at TP’s, making real-time diagnostics about the occurrence of TP’s a very difficult task. As a consequence, real-time level- and real-time TP-detection seem to be incongruent criteria.

A more general optimization criterion is needed which enables to control for both the amplitude and the time delay function. More precisely, the filter should damp high frequencies more heavily without increasing the time delay too markedly in the passband. For that purpose, consider a decomposition of (1.5) proposed in Wildi [3]

\[
\frac{2\pi}{T} \sum_{k=0}^{\lfloor \frac{T}{2} \rfloor} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 I_N(\omega_k) = \\
\frac{2\pi}{T} \sum_{k=0}^{\lfloor \frac{T}{2} \rfloor} |\Gamma(\omega_k) - \hat{A}(\omega_k)|^2 I_N(\omega_k) + \\
2 \frac{2\pi}{T} \sum_{k=0}^{\lfloor \frac{T}{2} \rfloor} A(\omega_k) \hat{A}(\omega_k) (1 - \cos(\hat{\Phi}(\omega_k))) I_N(\omega_k)
\]

where \( \hat{A}(\omega_k), \hat{\Phi}(\omega_k) \) denote amplitude and phase functions of the concurrent filter. The first summand on the right of the above equality corresponds to the fraction of the error variance which is attributable to imperfect selectivity properties (for example leakage-effect in the stop band) i.e. to the imperfect amplitude function of the concurrent filter whereas the second one measures the error variance attributable to its time delay. The above requirements – stronger damping of high frequency components and control of the time delay – may be obtained by the following generalized optimization criterion

\[
\text{min}_{\lambda} \left\{ \frac{2\pi}{T} \sum_{k=0}^{\lfloor \frac{T}{2} \rfloor} W(\omega_k)^2 |\Gamma(\omega_k) - \hat{A}(\omega_k)|^2 I_N(\omega_k) \\
+ \lambda \frac{2\pi}{T} \sum_{k=0}^{\lfloor \frac{T}{2} \rfloor} 2W(\omega_k)^2 A(\omega_k) \hat{A}(\omega_k) (1 - \cos(\hat{\Phi}(\omega_k))) I_N(\omega_k) \right\}
\]

where \( W(\omega_k) \) is monotonic in \( \omega_k \) (so that the effect of high frequency components is artificially magnified) and \( \lambda \geq 1 \) assigns more weight to the time delay error component. For \( W(\omega_k) = 1 \) and \( \lambda = 1 \) (1.7) reduces to (1.6) i.e. the best level filter is obtained. The trend filter of the newly released KOF-Konjunkturbarometer is based on (1.7).
Empirical comparisons: X-11 vs. DFA

The figures in section A.1 in the appendix compare real-time estimates of the DFA (based on (1.7)) and of X-11 for the most recent (last six) TP’s of the Konjunkturbarometer. When interpreting the results it is important to recall that the DFA was not explicitly optimized for the X-11 trend signal but for the ideal trend instead. Therefore, further improvements are obtained when referring to the latter signal, as shown below. Also, the filter is explicitly optimized for detecting TP’s at the costs of worse level approximations as can be seen in these figures.

In order to analyze important properties of the concurrent filters we here propose a general error-typology based on three particular TP’s in the following Figure 5.

- Error type 1:
  - ‘Consistent’ delay or anticipation around TP’s: as can be seen in the lower-right panel of Figure 5 (TP6) the last observable TP occurred in April 2005. However, the DFA detects it one month earlier, in March already, whereas X-11 does not reveal it until July i.e. one quarter later. Consistency here means that the real-time filter does not continuously revise its direction in the vicinity of the turning point.

- Error type 2. We distinguish two categories of errors of the second type:
  - ‘Inconsistent’ diagnostics at TP’s. As seen in the upper-left panel of Figure 5 (TP1) the X-11 TP occured in June 2000. The DFA detects it in May (error type 1: consistent anticipation) and X-11 indicates it the first time in May too. But in June, X-11 indicates a continuation of the preceding up-swing. Finally, the down-sing is consistently indicated from July on. In comparison to the DFA the diagnostics are inconsistent at the TP.

  - False diagnostics between TP’s. As seen in the upper-right panel of Figure 5 (TP3), the TP defined by the X-11 trend occurred in June 2002 whereas the DFA and X-11 detect it synchronously in September (consistent delay). However, as seen in the lower-left panel, the behaviour of X-11 between TP3 and TP4 is very complex, ranging from false unimodal to uninterpretable bimodal trend figures. Therefore, a correct appreciation of the business cycle is completely impossible in that period, which lasts more than one year.
Errors of the second type are due to leakage problems i.e. insufficient control of the amplitude function in the stop band whereas errors of the first type (especially delays) are due to insufficient phase control in the pass band of the boundary filters. Since favourable amplitude and phase characteristics of a concurrent filter are generally antagonistic (at least for the optimal filter), type 1 and type 2 errors are to some extent ‘antagonistic’ too. However, as shown in the tables below, the optimal DFA TP-filter clearly improves with respect to both error types which confirms that X-11 is an inefficient device. Note also that the level approximation by the DFA TP-filter isn’t optimal (as can be seen for TP3 and TP6 in the above figures), which confirms our previous statement that real-time level estimation and real-time TP-detection are two incongruent criteria.
Type 2 errors are especially problematic, because additional time is lost until a TP is confirmed, see strategy 2 below. The behaviour of X-11 between TP3 and TP4 in Figure 5 illustrates how long such ‘uncertainty’ periods may last. Therefore, different simple ‘objective’ strategies for identifying TP’s are evaluated in the following tables according to their impact on the number of type 1 and type 2 errors.

**Strategy 1**

The first identification strategy for TP’s is based on the first increment of the boundary filter only. More precisely, we count the number of times the sign of the first increment of the concurrent filter disagrees with the sign of the corresponding symmetric trend increment and discriminate the three error types accordingly. As an example, consider TP1 in the upper-left panel of Figure 5: the first increment of the concurrent X-11 filter, ending in May 2000, is negative (indicating a down-swing) whereas the corresponding trend increment is positive and the first increment of the following X-11 filter (ending in June 2000) is positive (indicating an up-swing) whereas the trend shows in the opposite direction. In the following table both errors are reported in the row “type 2 (at TP’s)” in columns “-1” (1 time unit before the TP occurred) and “0” (synchronous with the TP).

<table>
<thead>
<tr>
<th>X-11</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error type 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Type 2 (between TPs)</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 2 (at TPs)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total False diagnostics</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: X-11 trend and X-11 concurrent filter

From the figures in appendix A.1 we count five type 1 errors: three delays of two months at TP2, TP3 and TP6 (which are reported in column “2”) and two delays of three months at TP4 and TP5 (which are reported in column “3”). Moreover, three false diagnostics are located between TP2 and TP3. Note that by focusing on the first increment only in strategy 1, the very poor performance of X-11 between TP2 and TP3 can be alleviated to some extent. The total number of false diagnostics is obtained as 3*2+2*3+3+1+1=17.

The next table summarizes corresponding results for the DFA as shown in the figures in the appendix A.1: no errors of type 2, three anticipations (TP1, TP4 and TP6), one delay of one month (TP 5) and 2 delays of 2 months (TP2 and TP3). Again, the total false diagnostics are obtained as 3*1+1+2*2=8.

<table>
<thead>
<tr>
<th>DFA</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error type 1</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Type 2 (between TPs)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 2 (at TPs)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total false diagnostics</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
By discriminating the error types, we are able to quantify the mean delays of the procedures which are 2.4 for X-11 and 1/3 for the DFA. If we declare that type 1 anticipations are ‘beneficial’ to the intended purpose (which is to lead), then the total false diagnostics of the DFA shrinks from 8 to 5. Whichever case is considered, the DFA provides a huge improvement over real-time abilities of X-11.

Recall that the DFA has been optimized with respect to the ideal trend. Results for strategy 1 based on the ideal trend reference signal are summarized in the following table and graphical comparisons are provided in section A.2 in the appendix.

<table>
<thead>
<tr>
<th>Error type</th>
<th>DFA</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 2 (between TP's)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 2 (at TP's)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Total false diagnostics</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Two TP’s (TP5 and TP6) are lost towards the current boundary because the symmetric filter cannot be computed there. The mean time delay is accordingly \((-2+1)/(6-2)=-0.25\). If one ignores anticipations, then the DFA induces 1 false diagnostic only (a delay of one month at TP 2) for the observed four TP’s as defined by the ideal trend. As expected, the performance of the DFA TP-filter measured with respect to its own reference signal has significantly improved although the ideal trend is more difficult to approximate since it is smoother than the X-11 trend.

**Strategy 2**

Since X-11 is subject to type 2 errors we briefly analyze a second strategy for which a TP is detected as soon as two consecutive increments agree on their signs\(^6\). As an example, consider TP1: the first increment changes its sign in May 2000 but in June 2000 the confirmation is not realized. In August, the sign of July is confirmed, so strategy 2 would recognize the TP in August. As a consequence, the type 2 inconsistency has been avoided but the time delay has increased, generating a type 1 error. As can be seen by comparing table 1 with table 4, strategy 2 basically transforms type 2 into type 1 errors because the necessity of a confirmation increases the time delay of the filter.

---

\(^5\) Three two-month delays and two three-month delays realized in 5 TP’s (TP1 is another error type): so the mean delay of X-11 is \((3*2+2*3)/5=2.4\). For the DFA an analogous calculation gives \((-3+1+2*2)/6=1/3\) (since there no type 2 errors we have to take the mean over all 6 TP’s).

\(^6\) Waiting for a confirmation is a widely used practice, see for example Bry G., Boschan C. [1].
The total number of false diagnostics is obtained as (2+3*3+2*4+2)=21 and the mean time delay (type 1 errors) is (2+3*3+2*4)/6~3.2 i.e. the mean time delay has increased by approximately one time unit in comparison to strategy 1.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error type 1</td>
<td>1</td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Type 2 (between TP’s)</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 2 (at TP’s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total false diagnostics</strong></td>
<td><strong>21</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: X-11 trend and concurrent filter

The performance of real-time estimates of the well-known X-11 procedure and of the DFA have been compared specifically with respect to their ability of detecting TP’s of the KOF-Konjunkturbarometer. For that particular purpose, a generalization of the (traditional) level approximation criterion has been proposed. Performances have been computed with respect to the original X-11 trend as well as for the so-called ideal trend, for which the DFA has been explicitly optimized, and two simple identification strategies for TP’s have been proposed. For the time span including the most recent (last six) TP’s, diagnostics based on X-11 lead to 17 false diagnostics for the better strategy (strategy 1) whereas the DFA lead to only 8 false diagnostics (5 if anticipations are ignored). With respect to its own reference signal – the ideal trend – the DFA generated only 3 false diagnostics (1 if anticipations are ignored) in the same time span. The proposed concurrent TP-filter performs extremely well with respect to the intended purpose at costs of poorer level approximations which confirms that both criteria – level approximation and detection of TP’s – are incongruent. However, any `compromise´ can be easily realized in the above generalized framework.

**Summary**

The performance of real-time estimates of the well-known X-11 procedure and of the DFA have been compared specifically with respect to their ability of detecting TP’s of the KOF-Konjunkturbarometer. For that particular purpose, a generalization of the (traditional) level approximation criterion has been proposed. Performances have been computed with respect to the original X-11 trend as well as for the so-called ideal trend, for which the DFA has been explicitly optimized, and two simple identification strategies for TP’s have been proposed. For the time span including the most recent (last six) TP’s, diagnostics based on X-11 lead to 17 false diagnostics for the better strategy (strategy 1) whereas the DFA lead to only 8 false diagnostics (5 if anticipations are ignored). With respect to its own reference signal – the ideal trend – the DFA generated only 3 false diagnostics (1 if anticipations are ignored) in the same time span. The proposed concurrent TP-filter performs extremely well with respect to the intended purpose at costs of poorer level approximations which confirms that both criteria – level approximation and detection of TP’s – are incongruent. However, any `compromise´ can be easily realized in the above generalized framework.
Appendix

A.1 X-11 vs. DFA: X-11 reference signal
A.2 DFA concurrent filter, ideal and X-11 trend references

TP 1

TP 2

TP 3

TP 4

